Virtual Interface (VI) Architecture

Defining the Path to Low-Cost High-Performance Scalable Clusters

A Technology Imperative for Today's Business Climate

In the "old days" of the computing industry, the IT manager's worst nightmare was arriving for work in the morning to find the overnight runs still executing, simply because the computer didn't have enough capacity to finish the work. Essentially, the system had run out of headroom, and as a result, the company couldn't take orders for goods or services.

Today's CEO has very similar worries, but for even more critical reasons. There is no overnight. There is no luxury of downtime during which to perform backups. Most companies are now competing under rules of growing a 7x24 business and driving it over the Internet. With the Internet, if business ever had a level playing field, it's now.

The Internet is driving the globalization of all businesses, large and small. International banks and huge corporations are no longer the only entities operating around the clock. Small local companies have 'round-the-clock/'round-the-world opportunities at their doorstep thanks to the universal nature of the World Wide Web.

As the 8-to-5 workday becomes an artifact of a disappearing era, companies that hope to succeed in today's business environment need technologies that deliver the headroom they require. These technologies must provide not just high performance and reliability, but also a level of scalability that supports sustainable business growth. And, to ensure a reasonable cap on the soaring costs of business computing, any key enabling technology must be affordable and manageable.

Intel Corporation, in collaboration with other industry leaders, is lending its expertise to the development of a new proposed open specification that helps address these issues. The preliminary Virtual Interface (VI) Architecture specification defines an emerging high-speed cluster communication interface, bringing a powerful technology to the client/server computing environment. Specifically designed for low-latency, high-bandwidth message-passing between clusters of servers and workstations, the VI Architecture specification will enable businesses to build affordable, high-performance scalable clusters based on Intel-architecture systems.

Intel has also collaborated with Oracle Corporation to validate the benefits of the VI Architecture specification. This joint effort will optimize Oracle's new database on the Intel-architecture platform.

Bringing Scalability and Affordability to Clustering Technology

The driving force behind VI Architecture is the economic and competitive pressure to cluster computers. The concept of clustering a group of two or more interconnected systems or computers so they appear as a single system to clients has been around for at least a decade. Large corporations have relied on this type of configuration to deliver the enterprise-class scalability, reliability, availability and manageability they need for high-performance applications, particularly in data warehousing and on-line transaction processing (OLTP) implementations. Users and IT professionals alike appreciate the configuration flexibility and sophisticated failover advantages that clustering affords.

Businesses like these want a way to realize all of the advantages that can be gained by building clusters around common building blocks of standard high-volume (SHV) servers, shrink-wrapped applications and off-the-shelf operating systems.

The VI Architecture specification reflects a new model that addresses the rapid and industry-wide escalation of networking and application demands. Today, companies of all

sizes are deploying Internet and intranet services and distributed applications -- mission-critical applications requiring high availability and enterprise-class database and groupware solutions requiring high scalability.

A Technical Overview of VI Architecture

Specification development on VI Architecture dates back to January 1996, when Intel engineers began investigating the barriers to using SHV servers and commercial operating systems as the foundation for clusters. At the time, existing chipset technology supporting high-performance Intel processors had proved extremely successful in symmetric multiprocessing (SMP) configurations in a single server sharing memory. However, scaling beyond a four-way configuration for greater computing performance required a proprietary cluster or a proprietary chipset.

Researchers observed that the most significant roadblock to achieving high-speed cluster communications was the time required to send messages through the operating systems' legacy protocol stack. The solution to this problem lay not in modifying operating systems or addressing wire speeds but in developing an entirely new architecture designed to avoid the TCP, IP, UDP communication bottleneck that choked the ability to move data and instructions from one cluster element to another.

VI Architecture will be that solution.

Protocols have remained essentially unchanged despite tremendous improvements in physical bandwidth and several generations of processor improvements. The update with VI Architecture will effectively lower software overhead and allow the commoditization of message-passing.

The problem: As messages pass from one cluster node to another, the IP stack in the operating system copies the data at least twice -- once on the sending side and once on the receiving side.

The solution: VI Architecture eliminates the intermediate copies, freeing up CPU cycles, reducing latency and lowering the impact on bandwidth.

VI Architecture avoids the need to switch to the communication stacks and, thereby, avoids the context switch overhead.

VI Architecture enables a process to avoid interrupts under heavy workloads and enables interrupts only on wait-for-completion.

The VI Architecture specification defines a way to bypass these layers of protocol stacks. While in the traditional model all control and data passes through the OS kernel. With VI Architecture, control and setup go through the kernel but the data is transferred directly to/from the application to/from the network interface.

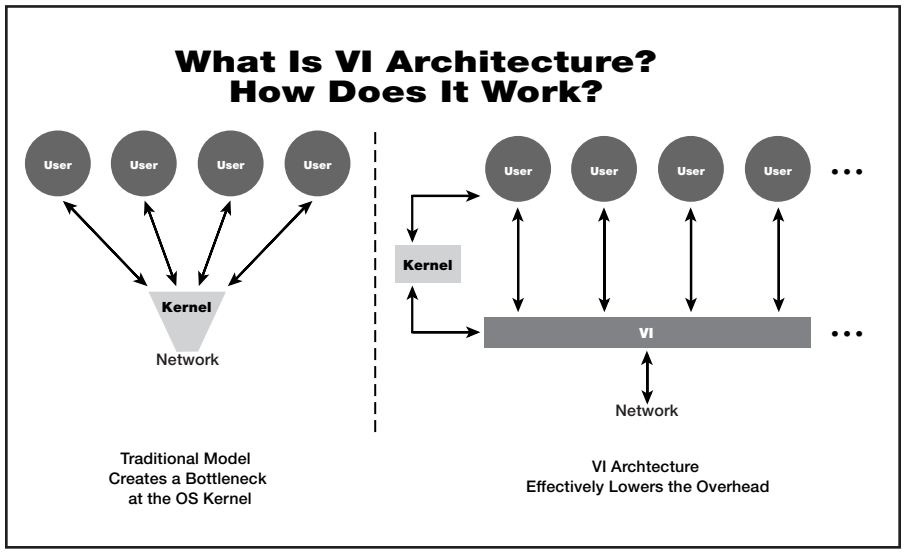

Fig. 1. Eliminating the bottleneck of the OS kernel not only boosts communication performance but also increases the number of CPU cycles available for performing other work.

Implementing the Virtual Interface Concept in the Communication Stack

VI Architecture bridges the cluster communication gap that, until now, has existed between the applications and operating system at the top of the communication stack and the high-performance volume hardware at the bottom of the stack. The VI Architecture specification encourages innovation above the interface by OSVs and ISVs.

The VI Architecture specification defines a standard software interface and a hardware interface model. The software interface is designed to support a variety of efficient programming models to simplify application development and ensure optimal performance. The hardware interface model is designed to be compatible with standard networks, including ATM, Ethernet and Fiber Channel, as well as specialized System Area Network (SAN) products. (A SAN is a network that connects the nodes within a single cluster. SANs are optimized to provide low-latency, high-bandwidth communication within the cluster.)

Beyond Theory: Intel and Oracle Demonstrate Proof of Concept

Like Intel, Oracle recognized that enterprise customers want the increased capacity, headroom and reliability levels that no single SMP system can deliver at a commodity price. Also like Intel, Oracle understood that responding to these customer demands called for a breakthrough architectural approach — one designed around standards-based, high-performance clustering. The VI Architecture specification promises to be that standard for the breakthrough.

Intel and Oracle have created a technology demonstration of Oracle Parallel Server (OPS) and VI Architecture to validate the benefits of a standardized, high-performance software/hardware interface for System Area Networks.

Oracle chose to lend its support to the VI Architecture effort because the architecture promises significant benefits including reduced complexity and improved performance, while still encouraging a diverse and competitive marketplace for interconnect technologies and products. The scalability inherent in the VI Architecture specification allows clusters to grow over time as companies desire to grow to accommodate larger transactions and/or more users for existing or new applications that access a common database.

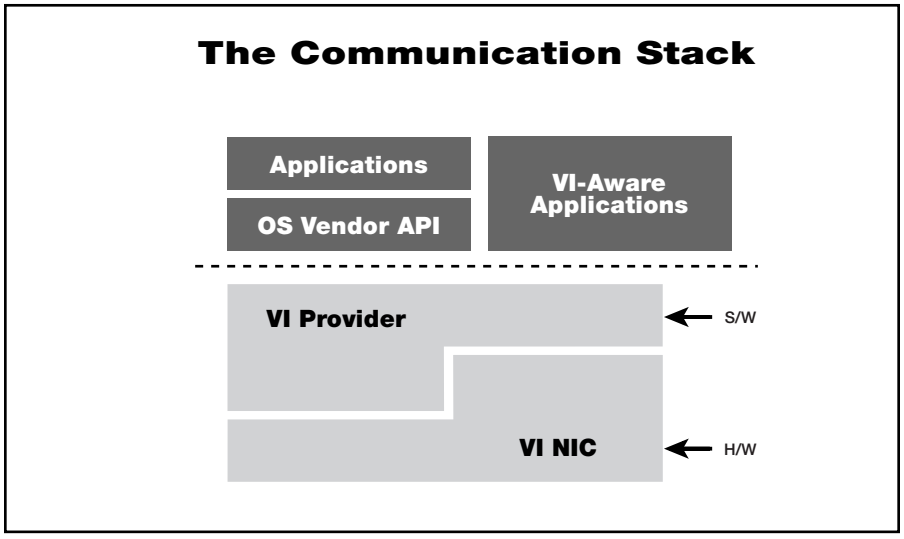

Fig. 2. The VI Provider Layer (VIPL) gives the illusion of dedicated access to a network, enabling applications to transfer data to the network interface without the overhead imposed by operating system intervention.

Summarizing the Benefits of the VI Architecture

VI Architecture allows users to capitalize on the cost savings, flexibility and versatility that are fundamental in the high-volume computing market segment. Because it defines a generic way for applications to run transparently over large-scale, high-volume clusters, VI Architecture provides benefits that span many areas of the computing industry.

Volume Technology - VI Architecture effectively brings volume technologies to clustering, addressing enterprise-class requirements. It blends the economies of scale essential in large-scale manufacturing with the computing power of high-performance processors, offering IT managers attractive price/performance in a multi-system solution.

Investment Protection Through Scalability - VI Architecture makes it possible to scale a cluster by adding more nodes, enabling a system to grow as business demands rise. VI Architecture scalability thus protects a company's investment in hardware and software because existing equipment and applications are compatible with future products. VI Architecture also enhances the portability of applications by isolating the application from the details of hardware and OS platforms.

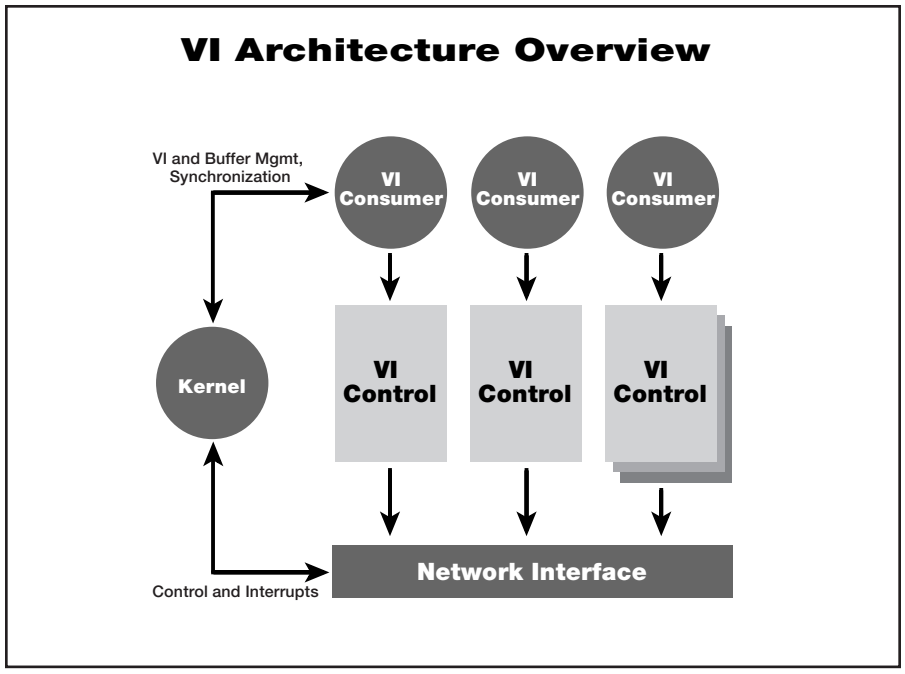

Fig. 3. VI Architecture is a fully protected, direct user-level access to the network interface.

Reliability - The VI Architecture specification takes a significant step toward defining SAN reliability. The VI Architecture specification guarantees applications a consistent level of reliability. In addition, it enables the application to choose higher degrees of reliability when provided by the underlying interconnect.

Benefits for Hardware Vendors - VI Architecture is expected to enable a new class of scalable cluster products offering high performance, reliability and improved availability. From a hardware product vendor's perspective, VI Architecture represents a framework for designing and building low-latency, high-reliability clusters of distributed message-passing systems.

Benefits for ISVs - Rather than worrying about the idiosyncrasies of various operating systems and hardware driver interfaces, software developers can program to the VI Architecture interface, which is consistent across platforms. This consistency makes it easy to design applications capable of running on a variety of scalable, cost-efficient hardware platforms. Besides the benefits that it offers in application portability, this consistency also simplifies the overall development process.

Galvanizing the Industry

VI Architecture momentum continues to build as OEMs, OS vendors, ISVs and IHVs embrace the industry initiative by developing products to the specification. The collective endeavor to define the specification, spearheaded by Intel in collaboration with Compaq Computer Corp. and Microsoft Corp. has grown to include more than 100 other industry leaders.

The promoters of the VI Architecture specification distributed a preliminary version of the specification to a long list of industry participants in January 1997. The final technical specification is expected to be released later in 1997. Native VI-compliant hardware implementations will be available in the first half of 1998.

 281895-001